Office de la propriété intellectuelle du Canada

Un organisme d'Industrie Canada

Canadian Intellectual Property Office

An Agency of Industry Canada

OIPE JC41 JAN 2 8 2002 PATENT & TRADEMARK OFFICE

RECEIVED JAN 3 1 2002 Technology Center 2100

Bureau canadien des brevets

Certification

Canadian Patent Office

Certification

La présente atteste que les documents ci-joints, dont la liste figure ci-dessous, sont des copies authentiques des documents déposés au Bureau des brevets.

This is to certify that the documents attached hereto and identified below are true copies of the documents on file in the Patent Office.

Specification and Drawings as originally filed, with Application for Patent Serial No: **2,354,301** on July 27, 2001, by **MATROX ELECTRONIC SYSTEMS LTD.,** assignee of Christian Simon and Djamel Yahia Meddah, for "Geometric Hashing for Model-Based Recognition of an Object".

RECEIVED FEB 0 4 2002 Technology Center 2600

S Grégoire

Agent certificateur/Certifying Officer

October 23, 2001

Date

Canada

(CIPO 68) 01-12-00

OPIC CIPO

9-10442-20CAPR

GEOMETRIC HASHING FOR MODEL-BASED RECOGNITION OF AN OBJECT

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This is the first application filed for the present invention.

MICROFICHE APPENDIX

[0002] Not Applicable

TECHNICAL FIELD

[0003] The present invention relates to machine vision, and in particular to geometric hashing for model-based recognition of an object.

BACKGROUND OF THE INVENTION

[0004] Techniques of visual object (and/or pattern) recognition are increasingly important in automated manufacturing, biomedical engineering, cartography and many other fields. Model-based recognition techniques typically must solve the problem of finding, in an image acquired by a camera, an occurrence of a previously defined model that has been affected by affine transformation. Affine transformations may be defined as transformations in which straight lines remain straight and parallelism is preserved. Angles however, may undergo changes and differential scale changes may be introduced.

[0005] Images, which are the projection of a three-dimensional world onto a plane are dependant on the position, orientation and the intrinsic properties of the camera which is acquiring the image. Image distortions might be introduced by different scale factors in the X and

Y directions. Perspective distortions might be introduced due to the optical axis of the camera's lens being at an oblique angle to the object plane. Distortion might also be introduced by optical imperfections of the camera's lens. Finally, distortions might appear because the object is not seated on a planar surface.

[0006] Known object recognition algorithms process acquired images to find an occurrence of a match between an image and a model that is subject to affine transformation. When images are distorted (e.g., due to perspective, lens distortion, etc.) finding a match with the model requires, from the matching algorithm, more than affine transformation capability.

[0007] Geometric hashing, as described in "Geometric hashing: A generalized and Efficient Model-based Recognition Scheme" (Y. Lamdan and H. J. Wolfson, Second International Conference on Computer Vision, Dec 1988, pp 238-249), and "Affine Invariant Model-Based Object Recognition" (Y. Lamdan, J. T. Schwartz, H. J. Wolfson, IEEE Transactions on Robotics and Automation, Vol. 6, No. 5, October 1990) has been proposed as a method of finding occurrences between an image and a model with affine transformation and partial occlusion.

[0008] In known geometric hashing methods, models of objects are represented by interest points. For each pair of interest points, a respective coordinate system is defined using the pair as a basis. The location of each of the other interest points can then be calculated within the respective coordinate system, to produce a representation of the interest points that are invariant to rotations and translations. For each coordinate system (basis), the

calculated coordinates of each interest point is then used as an index to reference a corresponding bin of a hash table, into which a reference to the interest point and the basis is inserted. The fully populated hash table is intended to provide a representation of the model that is independent of rotation and scale effects, and contains sufficient information to enable a match to be made even when an object is partially occluded.

[0009] As is well known in the art, object recognition commences by acquiring an image of the object (e.g., using a gray-scale digital camera), and processing the image to detect points of interest. As with the model, each pair of interest points is used as a basis for a respective coordinate system, within which the locations of each of other interest points are calculated. These calculated coordinates are used to access corresponding bins of the hash table. If an accessed bin contains a reference to a model interest point (and basis), then that model interest point (and basis) is accorded a vote. The basis that accumulates the largest significant number of votes is adopted as candidate, and extracted for further analysis.

[0010] According to Lamdan and Wolfson ("Geometric hashing: A generalized and Efficient Model-based Recognition Scheme", supra), this geometric hashing technique can deal with partially occluded complex objects. However, in practice, geometric hashing often fails heavily occluded objects, or objects having simple geometric shapes (e.g., triangles, rectangles, etc). This is because, in either of these cases, edge detection and analysis generally yields a small number of points of interest that pertain to the object in question. Thus if an image includes a heavily occluded object, or a partially occluded

simple object, the number of detected points of interest may be too low to permit the correct record (model basis) to accumulate a significant number of votes.

[0011] In addition, noise in an acquired image can produce errors in the computation of the coordinates of interest points, which may result in incorrect coordinate values being used to access the hash table. The problem of imprecision and computation errors can affect both points which define bases, and interest points that are used to vote. Since interest point coordinate values are a function of the chosen basis, errors due to imprecision in basis point and in interest point locations are accumulated. These problems are a significant disadvantages of geometric hashing, and are discussed in "On the Error Analysis of Geometric Hashing" (Lamdan, H. J. Wolfson, Proceedings IEEE Conference, Computer Vision and Pattern Recognition, pages 22-27, 1991) and "On the Sensitivity of Geometric Hashing" (W. E. Grimson, D. P. Huttenlocher, Technical Report A. I. Memo 1250, Artificial Intelligence Laboratory, Massachusetts Institute of Technology, 1990).

[0012] To avoid the above drawbacks, some improvements on geometric hashing have been proposed. In particular, instead of points of interest, the use of lines as affine-invariant features to represent an object in an image has been suggested (See "A probabilistic Approach to Geometric Hashing using Line Features", Frank Chee-Da Tsai, Technical Report No. 640, Robotics Research Laboratory, Courant Institute of Mathematical Sciences, June 1993). In comparison to discrete points, lines provide a more robust representation of an object, because errors due to noise do not accumulate as rapidly as they do in calculating

coordinates of discrete points. According to Tsai, lines can be used as the basis of respective coordinate systems, and all computations are made in lines space.

[0013] While this approach yields a more robust recognition system for complex object, it is still unable to reliably detect heavily occluded objects, or simple objects with partial occlusion.

[0014] The invention provides a method to improve geometric hashing and to solve the problems explained above.

[0015]

SUMMARY OF THE INVENTION

[0016]

BRIEF DESCRIPTION OF THE DRAWINGS

[0017] Further features and advantages of the present invention will become apparent from the following detailed description, taken in combination with the appended drawings, in which:

[0018] FIG. 1

[0019] FIG. 2

[0020] It will be noted that throughout the appended drawings, like features are identified by like reference numerals.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENT

[0021] The invention provides a geometric hashing technique for reliably finding one or more occurrences of a model in an acquired image

[0022] model definition

[0023] Step 1: Analyze image to detect interest points. These are preferably local features of an image (e.g., edge points, edge discontinuities, etc.), but may also include locations derived from image features (e.g., a centroid).

[0024] Step 2: Approximate edges with primitives. Each primitive is a line segment or curve segment that is mapped through a sub-set of the interest points. A "best fit" may be used. However, preferably, a "less than best fit" approach is used to ensure that continuity is preserved. For example, a set of straight line segments can be used to approximate a continuously curved edge. If each line segment is a best-fit, then adjacent line segments will frequently not intersect at their end points, so that continuity of the edge will be lost. Preferably, primitives are mapped such that, for a continuous edge, adjacent primitives intersect at adjoining end points.

[0025] In order to improve stability of recognition, bases are preferably selected according to any one or more of the following rules:

- use only highly "stable" bases (i.e., those associated with comparatively large primitives;
- allow a maximum of two bases for any one primitive;
- distribute basis origins as evenly as possible over entire image. This may include forcing a basis into a

region of an image where features are indistinct if other portions of the image have clear features favoring calculation of multiple bases;

- select basis orientations to approximately follow nearby primitives rather than angling sharply across image features.

[0026] Step 3: Represent primitives as invariants. In general, this involves defining a plurality of bases, and, for each base, defining at each primitive relative to that base. In order to improve recognition stability, each base is preferably calculated from a relationship between two or more primitives. For example the origin of a basis may be calculated as: an intersection point between lines extending from two primitives; a centroid of two or more such intersections; or, preferably, a centroid of two or more such intersections weighted by the length at least one of the primitives participating in each intersection. Similarly, the orientation of a basis may be calculated as: parallel to a primitive near the origin; an average of the orientations of each of the primitives used to calculate the origin; or a weighted average the of the orientations of each of the primitives used to calculate the origin, weighted by the lengths of each of the involved primitives.

[0027] If desired, the dimensions of the invariant representations of the primitives can be normalized to render the model scale independent. However, this increases the size of the hash table and complicates the problem of error tolerance. Accordingly, in preferred embodiments dimensions are not normalized and therefore absolute dimensions are mapped to the hash table. This renders the model scale dependent. However, scale can be

handled by any of a number of approaches including: multiple hash tables (one hash table per scale step); adjusting the size and/or number of error bars to guarantee that at least one scale step will result in a target primitive intersecting an error bar; or populating the hash table with references for model primitives in each of a plurality of scale steps, in which case each reference in the hash table is of the form (basis, scale, primitive).

[0028] Step 4: Sample primitives with an arbitrary granularity. The granularity chosen is based on a balance between speed and robustness. In general, a finer granularity (i.e. more samples) requires more processing time but yields improved tolerance to occlusion.

[0029] Step 5: for each Sample:

[0030] Step 5a: Map sample coordinate to an associated bin in a hash table.

[0031] Step 5b: Insert a reference to (basis, primitive) in the bin, and into one or more adjacent bins to create "error bars", the number adjacent bins used, and their orientation relative to the "main" bin are selected based on an expected image resolution error of the image processing system. In particular, based on the resolution of a camera used to acquire an image, the location of any point (or an edge) can only be determined within some error. Accordingly, if a target primitive lies within a predetermined error zone (or corridor) around a model primitive, then the target primitive should be considered to lie on the model primitive. The size of the error bars are selected to ensure that this occurs.

[0032] Step 6: Save hash table.

During Run Time

[0033] Step R1: Analyze target image to detect interest points. These are preferably local features of an image (e.g., edge points, edge discontinuities, etc.), but may also include locations derived from image features (e.g., a centroid).

[0034] Step R2: Approximate edges with target primitives. Each target primitive is a line segment or curve segment that is mapped through a sub-set of the interest points, in the same manner as described above.

[0035] Step R3: Represent target primitives as invariants, in the same manner as described above.

[0036] Step R4: Find end-points of target primitives and locate corresponding bins in the hash table.

[0037] Step R5: Locate bins crossed by an imaginary line extending between the "end-point bins". Any method of representing a line as discrete elements (e.g., pixels) may be used for this purpose. For example, a conventional Bresenham method, which is normally used to identify the pixels of a screen that need to be illuminated in order to display a line. In the present invention this method can conveniently be used to identify hash table bins instead of pixels.

[0038] Step R6: If any located bin includes a reference to a (basis, model primitive), then:

[0039] Step R6a: Calculate the length of an orthogonal projection of the target primitive onto the referenced model primitive. This length is a region of overlap

between the target and model, lying within the error corridor about the model primitive, as discussed above.

[0040] Step R6b: Accumulate the calculated length in an "bucket" associated with the identified basis. The accumulated length stored in the bucket may be referred to as the "target coverage".

[0041] Step R7: Repeat steps R4 through R6 for each target primitive within the target image.

[0042] Step R8: The total length of the model primitives (for any one basis) can be referred to as the "total coverage". The ratio of the accumulated coverage for any one basis to the "total coverage" of the model is the "relative coverage" of that basis. The basis with the highest relative coverage is selected as a candidate for further analysis and verification of a match in between the basis and the model.

[0043] The embodiment(s) of the invention described above is(are) intended to be exemplary only. The scope of the invention is therefore intended to be limited solely by the scope of the appended claims.

WE CLAIM:

1. A method of matching a pattern to a predetermined model represented by a hash table, the method comprising steps of:

   analyzing the pattern to identify interest points;

   processing the interest points to define one or more primitives approximating edges of the pattern;

   representing the target primitives as invariants relative to a selected basis;

   finding respective end-point bins of the hash table corresponding to each end-point of each target primitive;

   searching the hash table between each of the end-point bins to identifying a bin containing a reference to a (basis, primitive) of the model;

   calculating an orthogonal projection of the target primitive onto the referenced model primitive; and

   accumulating the calculation result in a bucket associated with the referenced basis.

2. A method of populating a hash table, the method comprising steps of:

   analyzing a model image to identify interest points;

   processing the interest points to define one or more primitives approximating edges of the model image;

   representing the model primitives as invariants relative to a selected basis;

sampling each model primitive at a predetermined granularity;

mapping the location of each sample point to a corresponding bin of the hash table;

inserting an reference to the (basis, primitive) of the model into the bin; and

inserting the reference into two or more adjacent bins of the hash table.

3. A method of defining a basis of local coordinate system of a plurality of primitives approximating an image, the method comprising steps of:

calculating an origin of the local coordinate system based on a selected relationship between two or more primitives; and

calculating an orientation of the local coordinate system based on an orientation of at least one of the two or more primitives.

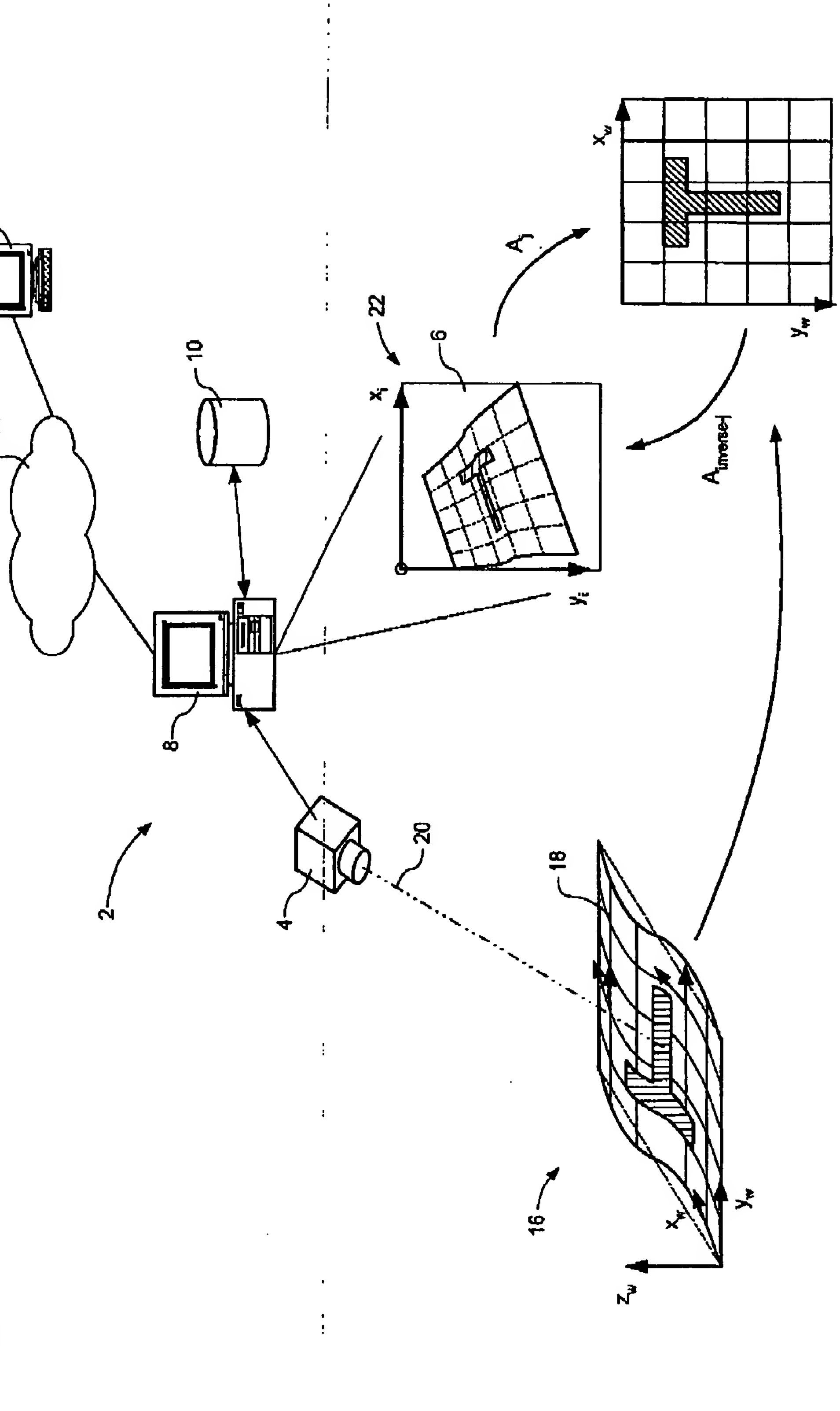

Figure 1

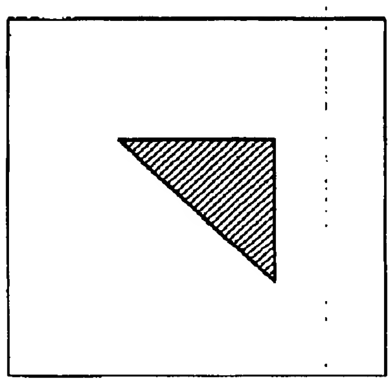

Figure 2a

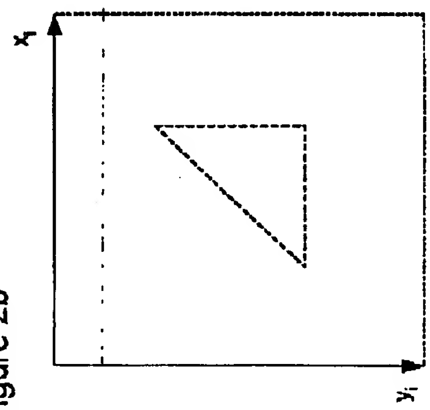

Figure 2b

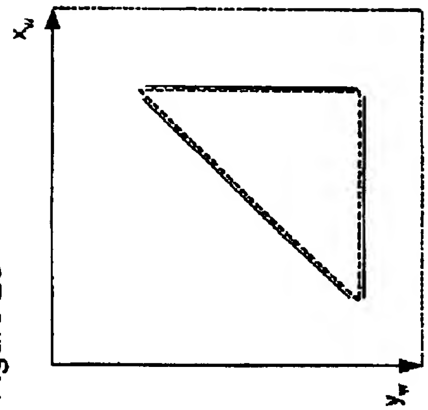

Figure 2c

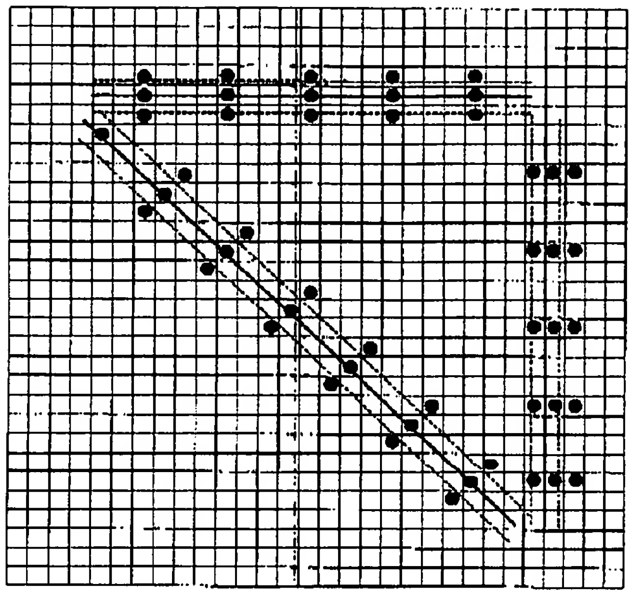

Figure 2e

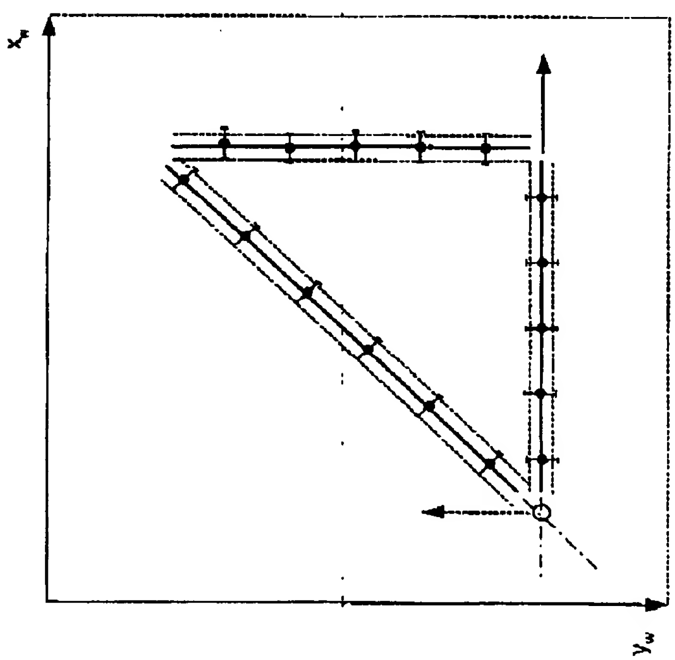

Figure 2d

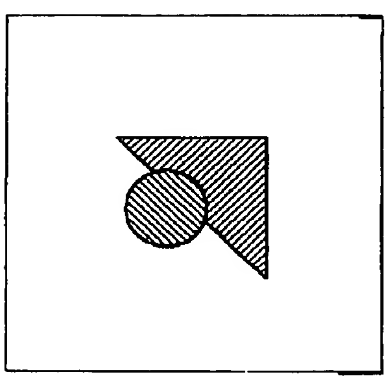

Figure 3a

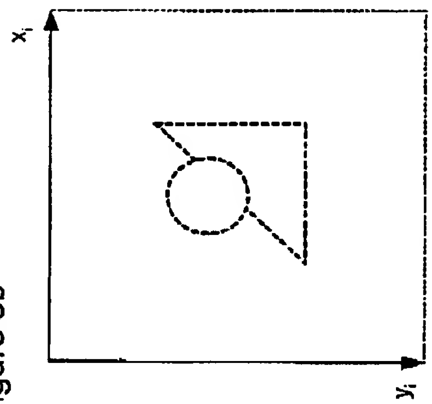

Figure 3b

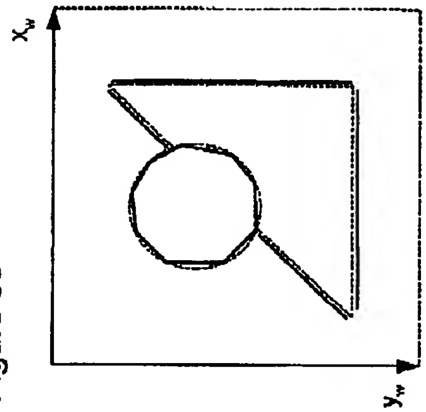

Figure 3c

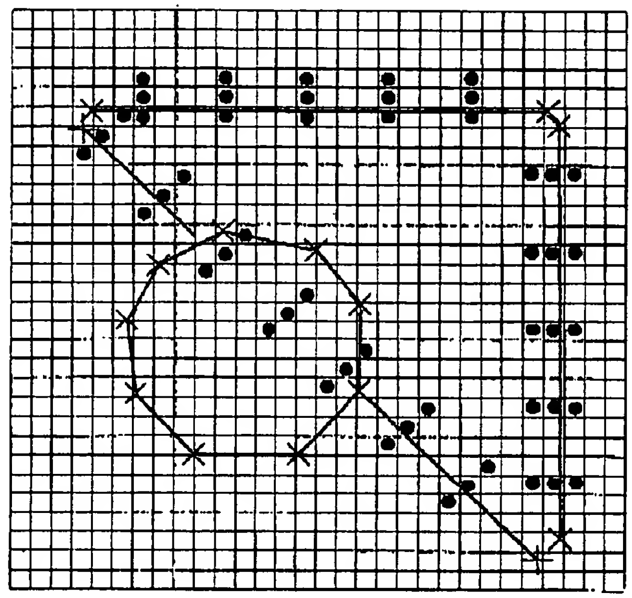

Figure 3e

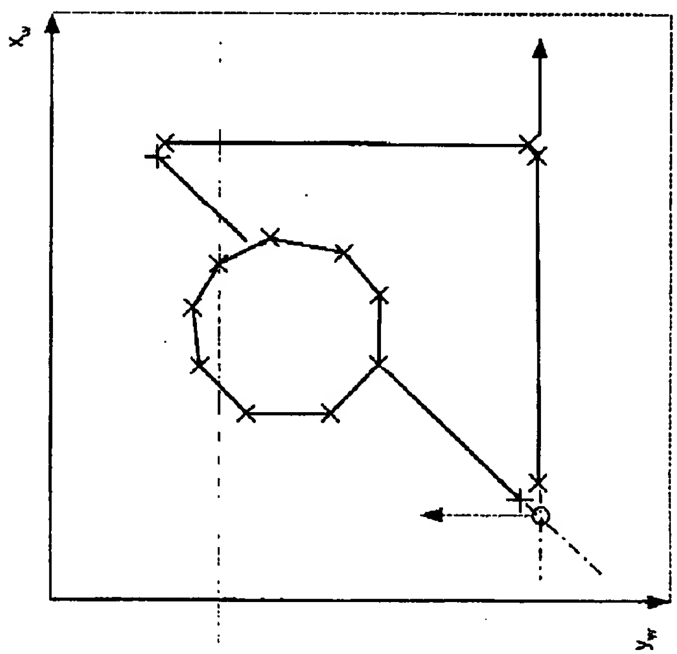

Figure 3d